



OPEN

## An electroencephalography connectome predictive model of major depressive disorder severity

Aya Kabbara<sup>1,2</sup>, Gabriel Robert<sup>3,4,5</sup>, Mohamad Khalil<sup>6,7</sup>, Marc Verin<sup>5,8</sup>, Pascal Benquet<sup>8</sup> & Mahmoud Hassan<sup>2,9</sup>✉

Emerging evidence showed that major depressive disorder (MDD) is associated with disruptions of brain structural and functional networks, rather than impairment of isolated brain region. Thus, connectome-based models capable of predicting the depression severity at the individual level can be clinically useful. Here, we applied a machine-learning approach to predict the severity of depression using resting-state networks derived from source-reconstructed Electroencephalography (EEG) signals. Using regression models and three independent EEG datasets (N = 328), we tested whether resting state functional connectivity could predict individual depression score. On the first dataset, results showed that individuals scores could be reasonably predicted ( $r = 0.6$ ,  $p = 4 \times 10^{-18}$ ) using intrinsic functional connectivity in the EEG alpha band (8–13 Hz). In particular, the brain regions which contributed the most to the predictive network belong to the default mode network. We further tested the predictive potential of the established model by conducting two external validations on (N1 = 53, N2 = 154). Results showed statistically significant correlations between the predicted and the measured depression scale scores ( $r_1 = 0.52$ ,  $r_2 = 0.44$ ,  $p < 0.001$ ). These findings lay the foundation for developing a generalizable and scientifically interpretable EEG network-based markers that can ultimately support clinicians in a biologically-based characterization of MDD.

Major Depressive Disorder (MDD) is one of the most common psychiatric disorders, mainly characterized by anhedonia and disturbed mood affecting the patient's quality of life and psychological state and has implications for their social and economic conditions<sup>1</sup>. Due to its increasing prevalence, chronicity, recurrence and degraded quality of life, MDD is now considered as a public health problem<sup>2,3</sup>. Currently, there are no biological signature of MDD, most probably because of its heterogeneity, and therefore prognosis (including better treatment response) and even diagnostic can be, sometimes, challenging. Researchers in mental illness advocate for a more biologically based framework to diagnose and treat these disorders, including depression<sup>4,5</sup>. At the cerebral level, most recent advances have moved from localized cerebral area disruptions to more network-based measures of mental disorders<sup>6</sup>.

Emerging evidence across functional studies consistently points at disruptions of MDD brain networks both in resting state<sup>7–19</sup> and during task-based connectivity<sup>20–22</sup>. Specifically, functional abnormalities in network topology<sup>11,20,21</sup>, modularity<sup>10,22</sup>, and efficiency<sup>8,11,15,16,23</sup> have been detected between MDD and healthy controls. are predominantly observed in the default-mode network (DMN) regions, including the posterior cingulate<sup>23,24</sup>, hippocampus<sup>7,8,23,25</sup>, parahippocampal gyrus<sup>7,21,23</sup>, precuneus<sup>23,26</sup> superior parietal lobule<sup>18,23,26</sup>, and the executive network (ECN) including dorsolateral prefrontal cortex<sup>8,9,14,25</sup> and the anterior cingulate cortex<sup>8–10,13,15–18,20,21,25</sup>. Importantly, DMN abnormal connectivity patterns is thought to be related with the cognitive vulnerability<sup>27,28</sup> and negative self-referential thoughts of MDD patients<sup>29,30</sup>.

<sup>1</sup>Lebanese Association for Scientific Research, Tripoli, Lebanon. <sup>2</sup>MINDig, F-35000 Rennes, France. <sup>3</sup>Academic Department of Psychiatry, Centre Hospitalier Guillaume R gnier, Rennes, France. <sup>4</sup>Empenn, U1228, IRISA, UMR 6074, Rennes, France. <sup>5</sup>Comportement et Noyaux Gris Centraux, EA 4712, CHU Rennes, Universit  de Rennes 1, 35000 Rennes, France. <sup>6</sup>Azm Center for Research in Biotechnology and Its Applications, EDST, Tripoli, Lebanon. <sup>7</sup>CRSI Research Center, Faculty of Engineering, Lebanese University, Beirut, Lebanon. <sup>8</sup>Univ Rennes, Inserm, LTSI-U1099, F-35000 Rennes, France. <sup>9</sup>School of Science and Engineering, Reykjavik University, Reykjavik, Iceland. ✉email: mahmoud.hassan.work@gmail.com

While this previous work has enabled a step toward a better understanding of the underlying pathophysiology of MDD, their translational potential into clinical use is hampered by the cost-effectiveness and availability of MRI. Electroencephalography (EEG) combined with the use of easy-to-implement measures of the brain network have been used to both binary classify<sup>31–33</sup> and to predict treatment outcomes<sup>34,35</sup> in MDD.

However, binary classification does not answer the question of MDD heterogeneity nor testing the possible continuum between normal sadness and pathological major depression. Very recently, resting state global connectivity measures predict both depressive and anxiety symptoms<sup>36–38</sup>. However, machine-learning/multivariate based connectivity measures are of high dimension and therefore often hampered by overfitting biases<sup>39</sup>. Proper analyses require large sample size and independent sample to validate the trained algorithm, which has not been the case in the field of EEG and affective disorders.

In this paper, we aim to establish a model that predicts depression severity based on resting-state EEG cortical networks. To tackle the problem of generalizability, three independent large datasets (N = 328) were used. Based on the prior findings, we hypothesized that the key regions that might contribute to the predictive model comprised DMN and ECN regions, mainly the anterior cingulate cortex, the prefrontal cortex, and the parahippocampal regions. We also expected that individual differences in MDD would be predicted by characteristics of alpha band as many studies suggest its relation to the pathologic characteristics of MDD<sup>40–42</sup>. EEG signals collected from 121 subjects (76 healthy controls and 45 MDD patients) were used to derive predictive models for depression severity. Brain networks were reconstructed using EEG source connectivity method<sup>43</sup>. Then, the Connectome-based Predictive Modeling (CPM)<sup>44</sup> was applied in order to establish the relationship between functional brain networks and depression severity in support vector regression model. The external validation was performed using two other independent EEG datasets, where the first includes 24 MDD and 29 healthy controls, and the second (N = 154) -with no MDD patients- shows an ultimate perspective of predicting a possible continuum between normal sadness and pathological major depression.

## Materials and methods

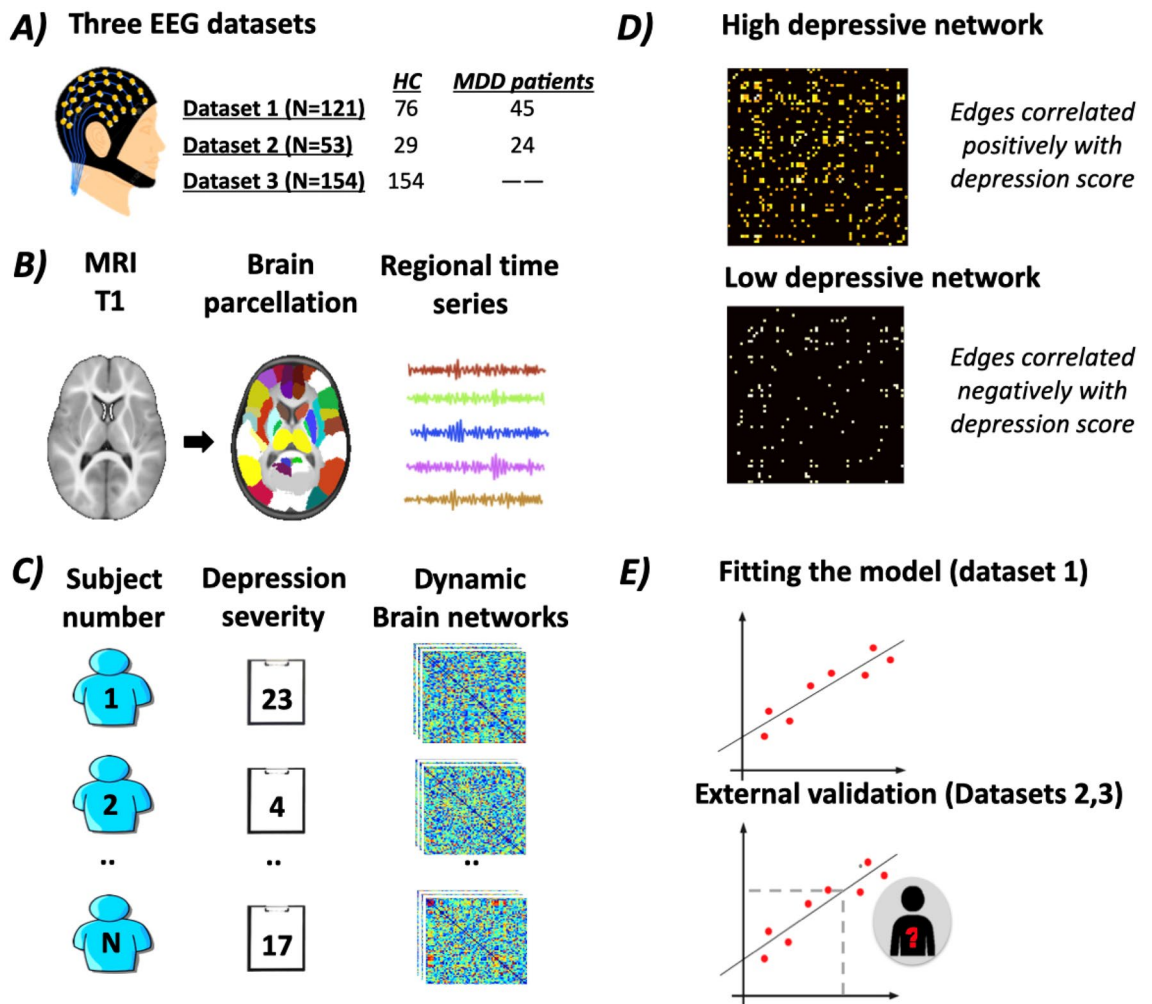
The full pipeline of this study is summarized in Fig. 1.

**First dataset. Participants.** Forty-five MDD patients and seventy-six healthy controls have participated in the current study. The EEG database is publicly available at <http://bit.ly/2rzY6ZY> and was used in a recent previous study<sup>45</sup>. The study was approved by the ethical committee of Arizona University and experiments were in accordance with relevant ethical guidelines. Prior to acquisition, all participants provided written informed consent. Participants were recruited from introductory psychology classes based on mass survey scores of the Beck Depression Inventory (BDI). Recruitment criteria included: (1) age 18–25, (2) no history of head trauma or seizures, and (3) no current psychoactive medication use. Control participants (N = 76) have stable low BDI (< 7) between mass survey and preliminary assessment, no self-reported history of MDD, and no self-reported symptoms indicating the possibility of an Axis 1 disorder as indicated by computerized self-report completion of the Electronic Mini International Neuropsychological Interview (eMINI: Medical Outcome Systems, Jacksonville, FL, USA). Depressed participants have a stable high BDI (> 13). Participant demographics are reported in Table 1.

**EEG acquisition and pre-processing.** Participants were asked to stay relaxed for five minutes while EEG signals were recorded. Signals were acquired using 64 Ag/AgCl EEG electrodes (Synamps system) positioned according to the standard 10–20 system montage, two electro-oculogram electrodes (EOG) for horizontal and vertical movements. Signals were sampled at 500 Hz, bandpass filtered between 0.5 and 100 Hz. All electrodes' impedances were kept below 10 k $\Omega$ .

Because of contamination of the EEG signals by various types of artifacts, EEG (pre)processing was applied following the same steps proposed in several previous studies dealing with EEG resting-state data<sup>46–48</sup>. These steps are summarized as follow: (1) identification of bad channels providing signals that are either completely flat or contaminated by movement artifacts. This was performed by visual inspection, complemented by the power spectral density, (2) interpolation of identified bad channels in Brainstorm by using neighboring electrodes within a 5-cm radius<sup>49</sup>. (3) segmentation into 40-s length epochs and only epochs with voltage fluctuations between + 100  $\mu$ V and – 100  $\mu$ V were retained. For each participant, four artifact-free epochs of 40-s lengths were selected. This epoch length guarantees a good compromise between the needed temporal resolution and the results reproducibility as demonstrated in<sup>48</sup>. Due to poor signal quality, EEGs from three HC subjects were excluded from the study.

**Second dataset. Participants.** A total of 24 subjects (30.88  $\pm$  10.37 years, female = 11) diagnosed with major depressive disorder, and 29 healthy controls (31.45  $\pm$  9.15 years, female = 9) were recruited for the study. Patients with MDD were recruited from Lanzhou University Second Hospital, Gansu, China, diagnosed and recommended by professional psychiatrists while healthy subjects were recruited by posters. The study was approved by the Ethics Committee of the Second Affiliated Hospital of Lanzhou University. All procedures performed in the study were in accordance with the ethical guideline of the national research committee of Lanzhou University. Written informed consent was obtained from all subjects before the experiment began. All MDD patients received a structured MiniInternational Neuropsychiatric Interview (MINI) that meets the diagnostic criteria for major depression of Diagnostic and Statistical Manual of Mental Disorders (DSM) based on the DSM. The dataset is publicly available in. Depression symptoms were identified with Hamilton Depression Scale (HDRS).



**Figure 1.** The full pipeline of our study. (A) EEG signals were acquired from three datasets: (1) N = 121 subjects (76 healthy controls and 45 MDD patients), (2) N = 53 subjects (29 healthy controls and 24 MDD patients) (3) N = 154 healthy subjects. (B) A template Magnetic Resonance Imaging (MRI) was segmented into 68 regions of interest (ROIs) by the means of Desikan Killiany atlas. Then, the regional time series of each subject were reconstructed using the weighted minimum norm estimate inverse solution (WMNE). (C) The inputs of the connectome-based predictive modeling (CPM) are the connectivity matrices and the depression score of each subject (BDI -i.e. dataset1-, Hamilton -i.e. dataset2, dataset3-). The brain networks of each subject were obtained by computing the phase locking value between the regional time series. (b) The training step: across all subjects, each edge is correlated to the BDI score. Then, the algorithm selects the most important edges using significance testing ( $p < 0.01$ ). Two matrices are thus resulted: the first corresponds to the network positively correlated with BDI and the second represents the network negatively correlated with BDI. (E) SVR predictive model is built to find the relationship between BDI and the sum of edge weights of the matrices obtained in the previous step. This model is then applied to predict the depression severity of subjects from the second and third dataset.

**EEG acquisition and pre-processing.** During experiment, participants were required to relax, stay awake and try avoiding movements while 5 min eye-closed resting state EEG were recorded in a sound-proof room. 128-channel HydroCel Geodesic Sensor Net (Electrical Geodesics Inc., Oregon Eugene, USA) was used and electrode signals were referenced to Cz. During acquisition, each electrode kept an impedance below 50 k $\Omega$ . EEG pre-processing was performed in the same way dealt with the first dataset, leading to four artifact-free epochs of 40-s lengths for each participant.

**Third dataset. Participants.** 154 young healthy participants (N = 154,  $25.1 \pm 3.1$  years, range 20–35 years, 45 female) were provided by the Mind-Brain-Body dataset<sup>50</sup> publicly available in [http://fcon\\_1000.projects.nitrc.org/indi/retro/MPI\\_LEMON.html](http://fcon_1000.projects.nitrc.org/indi/retro/MPI_LEMON.html). The study was carried out in accordance with the Declaration of Helsinki and the study protocol was approved by the ethics committee at the medical faculty of the University of Leipzig (reference number 154/13-ff). All included participants provided written informed consent before any data acquisition. The dataset originally included also 75 elderly subjects who were removed from the analysis to avoid

	HC	MDD	p-value
Cases (N)	76	45	NaN
Gender (M/F)	36/40	12/33	0.02
Age (years)	18.9 ± 0.6	19 ± 0.55	0.15
<b>Symptom scores</b>			
BDI	1.6 ± 0.75	22 ± 2.7	1E-20
BDI_AnH	0.16 ± 0.2	4 ± 0.8	1E-22
BDI_Mel	0.81 ± 0.4	6 ± 1	1E-17
TAI	30.87 ± 2.8	56 ± 4.6	1E-20

**Table 1.** Demographics of the 121 participants. *BDI\_AnH* anhedonia subscale of BDI, *BDI\_Mel* Melancholia subscale of BDI, *TAI* trait anxiety inventory.

the impact of age on the external validity of the model. Depression symptoms were identified with Hamilton Depression Scale<sup>51</sup>.

**EEG acquisition and pre-processing.** Participants were instructed to be awake with their eyes open and fixate on a low-contrast fixation cross on grey background. For each subject, 16-min resting state EEG was recorded with a ‘BrainAmp plus’ amplifier EEG using both 62-channel (61 scalp electrodes plus 1 electrode recording the VEOG below the right eye) and active ActiCAP electrodes (Brain Products GmbH, Gilching, Germany) positioned according to the international standard 10–20 extended localization system. Electrodes were referenced to FCz, the ground was located at the sternum and electrode impedance was kept below 5 kΩ. EEG were band-pass filtered between 0.015 Hz and 1 kHz and sampled at 2500 Hz.

Data were provided pre-processed, after passing through a pipeline that removed artefactual segments, identified faulty recording channels, and regressed out artefacts which appear as independent components in an Independent Component Analysis (ICA) decomposition with clear artefactual temporal signatures (such as eye blinks or cardiac interference).

**Brain networks construction.** Brain networks were reconstructed using the “EEG source connectivity” method<sup>43</sup> following two main steps: (i) estimate the cortical sources and reconstruct their temporal dynamics by solving the inverse problem, and (ii) measure the functional connectivity between the reconstructed regional time-series.

Brain networks were reconstructed using the “EEG source connectivity” method<sup>52</sup> using two main steps: (i) estimate the cortical sources and reconstruct their temporal dynamics using the inverse solution, and (ii) measuring the functional connectivity between the reconstructed regional time-series.

In brief, after co-registering EEGs and MRI template (ICBM152), a realistic head model was built using the OpenMEEG tool<sup>53</sup>. Then, the cortical surface was parcellated into 68 regions of interest by the means of Desikan-Killiany atlas<sup>54</sup>. As an inverse solution, the weighted minimum norm estimate (wMNE) algorithm was used to estimate the regional time series<sup>55</sup>. Afterwards, we filtered the reconstructed regional time series in the EEG frequency bands (delta: 1–4 Hz; theta: 4–8 Hz; alpha: 8–13 Hz; beta: 13–30 Hz and gamma: 30–45 Hz). In order to finally reconstruct the functional network, the Phase Locking Value (PLV) was assessed between the regional time series. This process applied yields an undirected and weighted connectivity matrix for each subject. We used the wMNE/PLV combination to reconstruct the dynamic networks, as it is widely used in the context of EEG source-space networks at rest<sup>48,56</sup> and it is supported by several model-based and real data-based comparative studies<sup>57</sup>.

**Connectome predictive modeling.** In this study, we are more interested in quantifying MDD heterogeneity in terms of brain connectivity rather than classifying binary groups. The binary classification will be very limited to deal with the MDD heterogeneity or testing the possible continuum between normal sadness and pathological major depression. Thus, we used the BDI scores, that quantify MDD severity, to construct the predictive networks of depression using the first dataset. This was performed using the connectome-based predictive modeling (CPM). Using the established model, external validation will be conducted using datasets 2 and 3. The CPM<sup>44</sup> is a recently developed method for identifying and modeling a brain network associated with a variable of interest, the depression severity in our case. CPM was previously employed in a number of studies to predict network alteration in several brain disorders such as anxiety related illnesses<sup>58</sup> and sleep disorders<sup>59</sup> as well in some other conditions such as personality traits and creativity<sup>60–62</sup>.

**Data preparation.** Before establishing the predictive model, we prepared the inputs required for the prediction:

- 1- Preparation of the resting state functional connectivity matrices: For each subject, the connectivity matrices were averaged across time samples and epochs in order to obtain one representative matrix.
- 2- Preparation of the depression severity scores: Due to the difference in the measure used to assess the depression severity across datasets and their range of values (BDI scores (0–63) in Dataset 1 and HDRS in Datasets 2



and 3 (0–52)), it is recommended to use standardized measures in order to improve the interpretability of the results<sup>63,64</sup>. More specifically, the standardized BDI ( $BDI_s$ ) scores of dataset 1 were calculated by computing the difference between BDI and the distribution-mean divided by the standard deviation (Z-transformation). To ensure that the same scale was used across datasets, Z-transformation of the Hamilton scores corresponding to Datasets 2 and 3 was performed using the equivalent population-mean and standard deviation calculated on Dataset 1, yielding to standardized Hamilton scores ( $HDRS_s$ ). More precisely, we have first calculated the mean and the standard deviation of the BDI scores of the training set. Then, we have computed the equivalent mean and standard deviation following the so-called rule of three mathematical rule based on the HDRS range (0–52). Hence, the mean and the variance used in the Z-transformation of HDRS scores are obtained such that:

$$\text{Equivalent\_metric\_HDRS} = \text{metric\_BDI} \times \frac{52}{63}$$

where metric is either the mean or the standard deviation,  $\text{Equivalent\_metric\_HDRS}$  is the equivalent value of the considered metric used in the normalization of Hamilton scores,  $\text{metric\_BDI}$  is the value of the considered metric as calculated based on BDI distribution. Finally,  $HDRS_s$  was calculated by subtracting the  $\text{equivalent\_mean\_HDRS}$  and dividing by the  $\text{equivalent\_standard\_deviation\_HDRS}$ .

**Internal validation: prediction analysis using cross-validation.** We established the predictive model using the first EEG dataset. To quantitatively assess the ability of the model in predicting novel samples, an internal validation using dataset 1 was performed. Thus, a fivefold cross validation was used to evaluate the prediction performance. Particularly, it was reported that K-fold cross-validation procedure provides an accurate model evaluation compared to holdout method and leave-one-out procedure<sup>65</sup>. Here, fivefold cross-validation method was performed by splitting the 121 subjects into 5 non-overlapping subsets of data (also known as folds). The assignment of subsets as training and testing data is repeated 5 times, where at each time, the subjects of 4 subsets are assigned as training data, with the remaining subset reserved as testing data. The training procedure includes the following steps:

1. *Edge-standardization* The weight of each edge was standardized (mean-centered) across the training set. Specially, we first computed the mean and standard deviation on each edge across subjects in the training set. Then, for each subject, we performed a Z-transformation on each edge by calculating the difference between its weight and the mean divided by the standard deviation. This step was performed in many previous CPM studies<sup>61,62,66</sup>.
2. *Identification of the predictive edges* Each edge in the connectivity matrix was correlated with the  $BDI_s$  scores for the training set using Pearson's correlation. To calculate this correlation, the Z-transformed weight of the edge was used. Afterwards, only edges that show significant correlations with  $BDI_s$  ( $p_{FDR} < 0.01$ ) were retained. In particular, the false discovery rate (FDR) approach was used to account for the multiple comparison problem. The correlation values were separated into a positive tail (i.e. edges correlated positively with  $BDI_s$  scores) and a negative tail (i.e. edges correlated negatively with  $BDI_s$  scores). Therefore, this step will result in the reconstruction of two networks: high depressive network (including connections in the positive tail) and low depressive network (including connections in the negative tail).
3. *Computation of the edge strength* The edge strengths were calculated in both positive and negative tails in order to correlate it with  $BDI_s$ , as proposed by<sup>44</sup>. By combining both depressive networks, we also calculated the summed index. This latter is obtained by summing the Z-scores of all connections of the positive and subtracting those obtained from the negative tails.
4. *Construction of Support Vector Regression model (SVR)* SVR was employed to respectively relate positive and negative edge strengths to  $BDI_s$  scores<sup>67</sup>.

Once the trained model is established using the training subset, the model is tested using the training fold. The testing procedure includes the following steps:

1. *Edge standardization* We have first standardized the edges of the depressive networks using the same parameters (i.e. mean and standard deviation) acquired during the training procedure.
2. *Testing the model* The trained SVR model was used to predict the testing participant's BDI scores.
3. *Assessing the predictive performance of the model* To evaluate the results of the prediction, we used several metrics: mean absolute error (MAE) which measures the difference between the predicted and the observed score, Pearson's correlation (R) that measures the correlation between the predicted and the observed score, and the coefficient of determination, (R-squared) that reports the proportion of the variation in the dependent variable (depressive score) that is predictable from the independent variable (edge strength).

The fivefold cross validation was repeated for 100 iterations, and the resulted performance presents the average performance across all iterations and folds.

Finally, in order to check if the obtained correlation is significantly better than expected by chance, we performed a permutation test by permuting the BDI scores of all participants for 1000 times. In each permutation, we shuffled the  $BDI_s$  scores in a random way and then we repeated the cross-validation procedure explained above including all the steps of the training and testing procedures. An averaged prediction error quantified by the mean absolute error (MAE) value across folds was obtained in each permutation. This yields in a null distribution of 1000 correlation values. Then, p-value was estimated after dividing the number of times the shuffled value was

greater than the true MAE value by 1000 (the number of iterations). Furthermore, the prediction results were validated using a different cross-validation scheme using  $K=10$ , to see the consistency of results.

**The predictive networks.** Due to the difference in the predictive networks obtained across iterations during the internal validation, the final generalized positive and negative predictive edges were defined as those that persist in all iterations. The derived common positive and negative networks will thus be used as the predictive networks for the external validation. The same procedure was applied in the previous connectome-based predictive modeling studies<sup>58,66</sup>.

**External validation.** The derived positive and negative predictive networks obtained using the first dataset were used as predictive networks for the second and third independent dataset. Before computing the edge strength within the high and low depressive networks, edge z-transformation was conducted using the standardization parameters obtained on Dataset 1. Then, the SVR trained models using the predictive edge strengths were used for testing. The predictive performance was assessed by calculating the mean absolute error (MAE), Pearson's correlation (R), the coefficient of determination (R-squared)<sup>63,65</sup> and bootstrapping  $p$ -value.

**Confounder analysis.** To explore potential confounding variables, we correlated age with the dependent variable (i.e. the depression severity) as well as with the independent variables that are the predictive selected features (i.e. the significant edges) using Pearson's correlation. Similarly, Point-Biserial Correlation was used to correlate gender with the depression severity and the predictive edges. To avoid potential confounder effects of correlated variables, partial Pearson correlation analysis were conducted with the considered variable in edge selection as suggested by<sup>58</sup>.

## Results

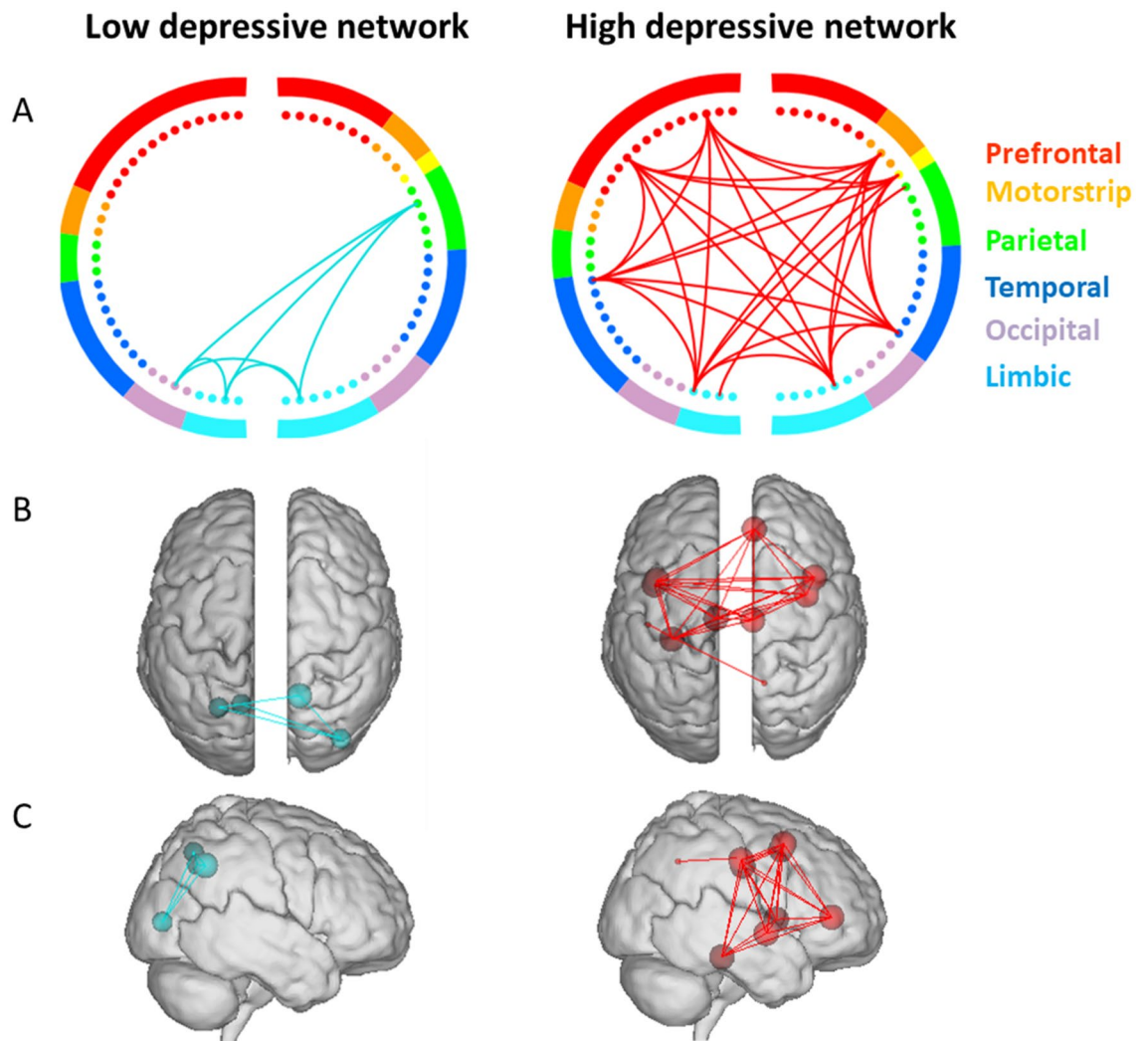
**Demographics data.** In dataset 1, the mean BDI is 9.5207 with a standard deviation of 10.5064. These values were used to obtain  $BDI_s$  that were the target variables in the model training. The equivalent values of the same parameters based on the Hamilton range were used to calculate the standardized Hamilton scores ( $HDRS_s$ ) that represent the target variables in external validation.  $BDI_s$  range from  $-0.9$  to  $1.94$  with a robust range (range between the 5th and 95th percentiles) of  $1.73$  (arbitrary units).  $HDRS_s$  in dataset 2 range from  $-0.9062$  to  $4.5245$  with an interquartile of  $3.707$ , while  $HDRS_s$  in dataset 3 range from  $-0.9062$  to  $0.7082$  with an interquartile of  $0.4613$ .

For dataset 1, no statistical difference in age ( $p=0.15$ ) was depicted between MDD and HC. Results show the existence of a statistical difference in gender ( $p=0.02$ ) as computed using chi-square test. MDD patients and HC differ in all the recorded psychiatric scores ( $p<0.0001$ ) as computed using the Wilcoxon ranksum test. For dataset 2, there were no significant between-group differences in age ( $p=0.832$ ), or gender ( $p=0.269$ ). Statistical differences in the self-reported PHQ-9, and the Generalized Anxiety Disorder-7 (GAD7), the Hamilton depressive scores were obtained ( $p<0.001$ ).

**Confounder control.** Results showed that age was not significantly associated with BDIs nor the weights of any predictive edge derived in the training dataset ( $p>0.2$ ). Ultimately, age was not considered to be a confounding variable in our predictive model. While gender was not significantly correlated with  $BDI_s$  ( $p>0.05$ ), its effect size is found to be considerable ( $r=0.15$ ,  $p=0.09$ ). Thus, gender was considered as a confounding variable in the model estimation.

**Predictive networks.** As mentioned in the materials section, the predictive networks were derived after applying the cross-validation procedure on dataset 1. Two main networks result: (1) The high depressive network which denotes the network in which edges show positive significant correlations with  $BDI_s$  scores, and (2) the low depressive network which denotes the network in which edges show negative significant correlations with  $BDI$  scores. Gender was considered as a covariate in the model estimation. No significant networks were detected in delta, theta, beta and gamma bands. In alpha band, the numbers of edges that contributed to the prediction ranged from 75 to 189 across all iterations. Notably, 38 of these edges (33 high depressive edges, 5 low depressive edges) appeared in every iteration and were defined as the depressive edges. The high depressive network exhibited dense functional connections in predominantly prefrontal, insula and limbic lobes (Fig. 2). More specifically, the implicated brain regions are: the caudal middle frontal gyrus (cMFG), insula (INS), parahippocampal (paraH), posterior cingulate (PCC), and rostral anterior cingulate (rACC). The low-depressive network showed edges within occipital, parietal and limbic lobes. Nodes are the lateral occipital gyrus (LOG), the superior parietal lobule (SPL) and the precuneus (PCUN). Notably, the implicated regions in the predictive networks belong mainly to the default mode network (paraH, PCUN, PCC, SPL).

**Internal validation using dataset 1: cross validation.** As the predictive performance differs across folds/iterations, we reported here the average performance metrics along with a measure of variability across folds/iterations (i.e. the standard deviation). Results showed that the model designed based on the predictive edges of the low depressive network (showing negative correlation with BDIs) were able to predict BDIs scores in novel individuals ( $r=0.61\pm 0.08$ ;  $MAE=0.68\pm 0.09$ ,  $R\text{-squared}=0.48\pm 0.08$ ,  $p=0.001$ ). Figure S3 shows the corresponding histogram of the null MAE distribution. According the high depressive network, the predictions obtained from of the model achieved  $r=0.42\pm 0.09$ ,  $MAE=0.94\pm 0.06$ ,  $R\text{-squared}=0.23\pm 0.05$  with  $p=0.003$  (see Fig. S2). Figure 3A,B illustrate an example of the internal validation obtained for the low and high depres-



**Figure 2.** Depictions of the high- and low-depressive networks. Circle plots (A), glass brains—top view (B), glass brain- right view were obtained by keeping the significant edges ( $p < 0.01$ , FDR corrected). Colors within the circle plots correspond to lobes of the brain.

sive networks, respectively. The illustrated examples are picked up from a single representative iteration chosen randomly on condition that the corresponding MAE value resides in the interquartile of the MAE distribution across iterations.

We also evaluated the predictions of the model that combined both high and low depressive edges. Our findings reveal a positive significant correlation of  $r = 0.59 \pm 0.1$ , between the observed and the predicted BDI<sub>s</sub>, with R-squared =  $0.46 \pm 0.07$ , MAE =  $0.75 \pm 0.091$  with  $p = 0.001$  (Fig. S1). An example is shown in Fig. 3C.

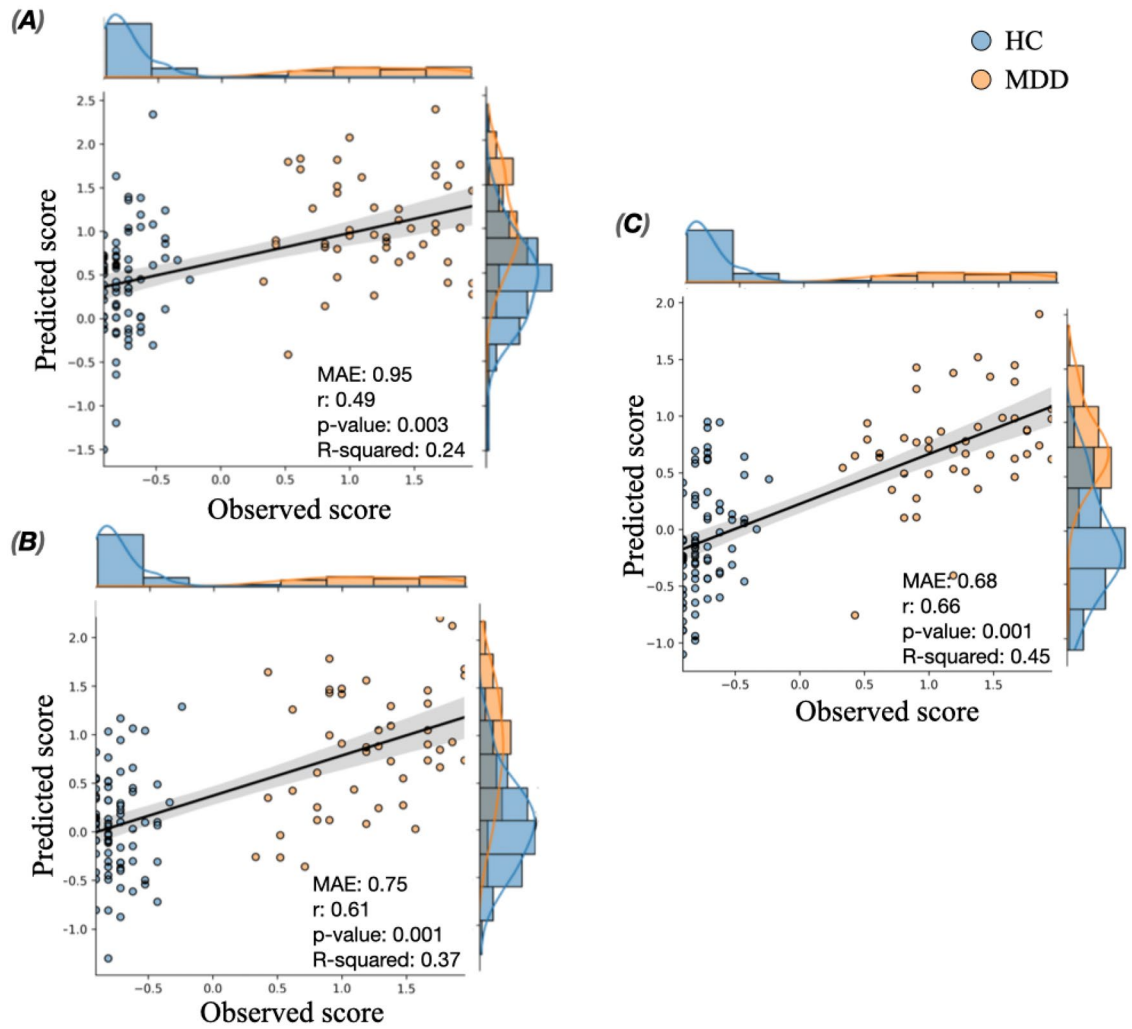
Results obtained using tenfold cross validation show consistent results (see Table 2).

**External validation on datasets 2 and 3.** Using the model developed based on the high depressive edges, the performance of the prediction achieved ( $r = 0.3965$ , MAE = 1.5984, R-squared = 0.1572,  $p = 0.002$  (Fig. S4, Fig. 4A). In addition, statistically significant prediction was obtained using the low depressive connections ( $r = 0.3198$ , MAE = 2.1268, R-squared = 0.1023,  $p = 0.02$  (Fig. S5, Fig. 4B). Using both predictive networks, the prediction achieved  $r = 0.5164$ , MAE = 1.4230, R-squared = 0.2666 with  $p = 0.001$  (Fig. S6, Fig. 4C).

Our findings reveal that the SVR model was able to predict the depression scores for the high-depressive network ( $r = 0.4391$ ; MAE = 0.4700, R-squared = 0.1928,  $p = 0.03$ ; Fig. 4D, Fig. S7), but not for the low-depressive network ( $r = 0.0881$ , MAE = 1.2980, R-squared = 0.0078,  $p = 0.45$ ; Fig. S8, Fig. 4E).

## Discussion

The objective of this study is to establish a connectome-based model capable of predicting depression score at the individual level. Using three independent resting state EEG datasets, our findings proved that depression could be predicted using intrinsic functional connectivity. More specifically, the majority of the nodes that contributed the most to the predictive network belong to the DMN obtained in the alpha frequency band. Interestingly,



**Figure 3.** A typical example (chosen from a random iteration) showing the relationship between the observed and the predicted scores using (A) low depressive edges, (B) high depressive edges, (C) combined depressive edges.

	Fivefold cross validation		Tenfold cross validation	
	High depressive edges	Low depressive edges	High depressive edges	Low depressive edges
r	0.42 ± 0.09	0.61 ± 0.08	0.47 ± 0.13	0.65 ± 0.11
MAE	0.95 ± 0.06	0.48 ± 0.09	0.91 ± 0.1	0.42 ± 0.09
Permutation-based p	0.009 ± 0.006	0.005 ± 0.003	0.006 ± 0.009	0.004 ± 0.006
R-squared	0.23 ± 0.05	0.48 ± 0.08	0.21 ± 0.05	0.43 ± 0.05

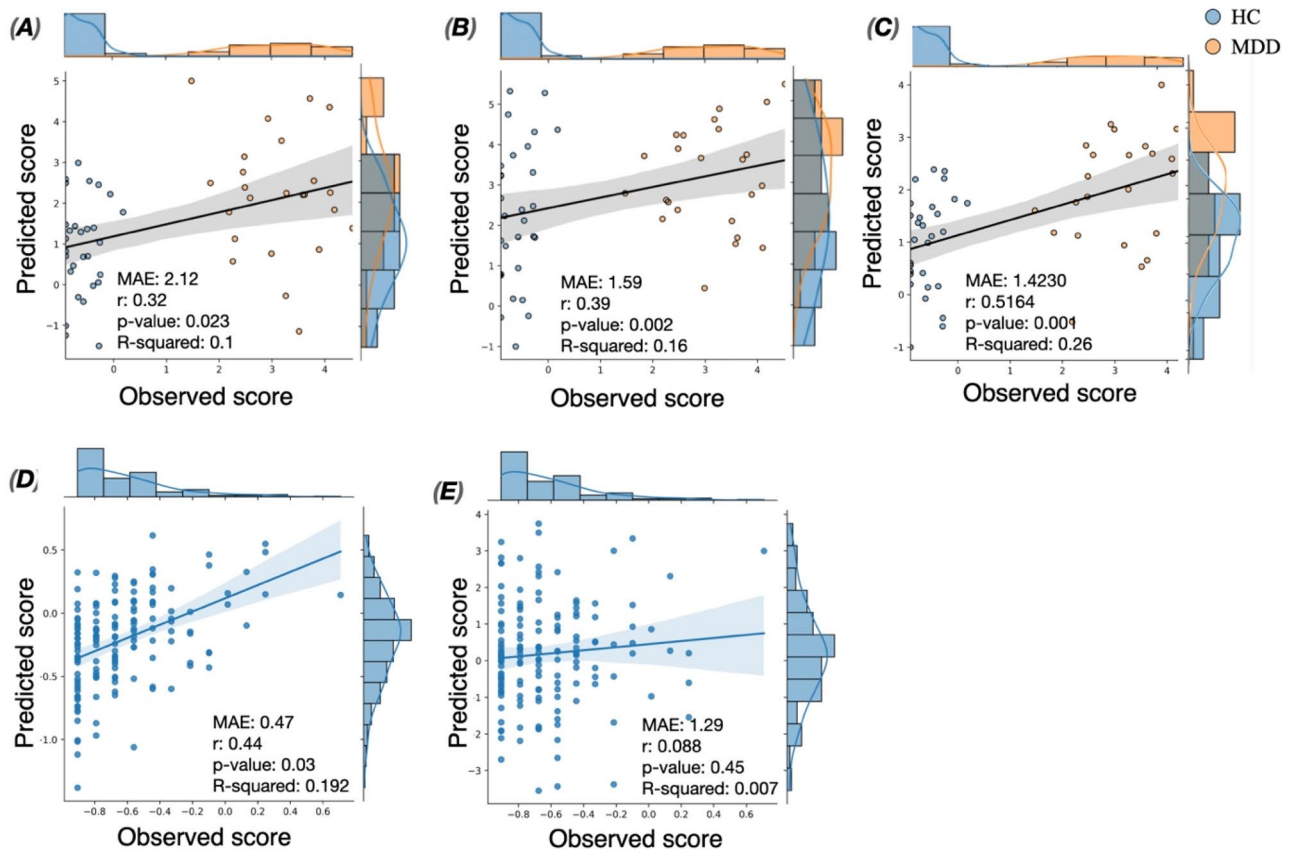
**Table 2.** Results of the internal validation using fivefold and tenfold cross validations. r denotes the Pearson’s correlation between observed and predicted scores. MAE denotes the mean absolute error of the prediction. R-squared presents the coefficient of determination.

generalizability of the model to predict novel and independent samples was demonstrated across different datasets as well as across different psychometric tests (BDI, HDRS).

The model predicts a considerable amount of the variance in individual depression symptom severity (48% with internal validation and 37%–19% with external validation). The obtained MAE obtained in internal and external validations suggest a good generalizability of the results. Results are discussed hereafter.

**Brain regions implicated in predictive network.** Alterations in the paraH connectivity were previously reported<sup>7,21,23</sup>. As this region plays a major role in memory retrieval, it is supposed that such alteration is associated with the recall or the imagination of negative memories and events in MDD patients<sup>68</sup>. Consistent





**Figure 4.** (A) Correlation between the predicted and observed Hamilton scaling score of the second dataset using high-depressive network, (B) the prediction results of the second dataset using low-depressive network, (C) the prediction results of the second dataset using both high and low depressive network. (D) The prediction results of the third dataset using high depressive network. (E) The prediction results of the third dataset using low depressive network.

with our findings, previous studies have reported higher activation of paraH in depressed subjects compared to healthy group during rumination (moments)<sup>69</sup>.

In addition, increased blood flow in left and right PCC regions were associated with MDD<sup>70</sup>. Using EEG source connectivity combined with graph theory, a recent EEG study shows that MDD brain network is characterized by increased efficiency in PCC<sup>23</sup>. Importantly, PCC is a key region in the limbic system and the DMN, and has a major role in regulating emotions and motivational thoughts<sup>71</sup>. Similarly, ACC which has been associated with MDD in multiple studies<sup>8-10,13,15-18,20,21,25</sup>, has been most frequently linked to the experience of pain, and specially in monitoring painful social situations such as exclusion or rejection<sup>72</sup>. This may explain the implication of PCC and ACC in the high depressive network. Furthermore, the prefrontal and limbic cortical regions are directly modulated by the subcortical structures that are involved in the negative emotional circuits. For instance, the canonical amygdala, hippocampal structures, and the lateral habenula (LHb) has been found to represent key brain regions in the pathophysiology of depression<sup>73</sup>. Clinical and preclinical evidence implicates hyperexcitability of the lateral habenula (LHb) in the development of psychiatric disorders including<sup>74</sup>. Hyperactive LHb strongly inhibit the dopaminergic motivational, cognitive and reward system<sup>75</sup> and produces depressive- and anxiety-like behaviors, anhedonia and aversion. Dopaminergic and anxiety circuits strongly modulate the limbic cortex, including the orbitofrontal cortex and the anterior cingulate cortex. The medial prefrontal cortex in rodent, the homologue of the anterior cingulate cortex of human are strongly involved in the expression of anxiety and are involved in MDD<sup>76</sup>. Because these internal emotional circuits strongly modulate the prefrontal, insula and limbic lobes, it is not surprising that in the high depressive network we found a dense functional connection between these areas.

According to the low depressive network, the regions implicated are SPL, LOG and PCUN. In fact, SPL, LOG and PCUN connectivity were found to be different between MDD patients and healthy controls in many studies<sup>18,23,26</sup>. SPL is principally involved in attention and visuomotor integration<sup>77</sup>, whilst LOG plays a major role in the visual perception. In addition, visuospatial processing, reflections upon self, and aspects of consciousness are associated with PCUN. This may explain the implication of these regions in low depressive subjects where attentional and visual performances are more efficient.

**EEG frequency bands in depression.** In this study, differences between MDD and healthy groups were only depicted in alpha band, and no significant results were obtained in other EEG frequency bands. Interestingly, using the same frequency band, the model has demonstrated its performance in predicting novel and independent samples. This finding is in line with a number of resting state EEG studies that continuously report MDD alterations in alpha frequency patterns. Particularly, alpha-asymmetry was suggested to be a potential marker for depression<sup>26,40,78,79</sup> whereas alterations in other frequency band such as delta, beta and gamma power were inconsistent<sup>40</sup>. More importantly, a connectivity-based EEG analysis suggests an increase in alpha-band connectivity between the anterior cingulate cortex and both the prefrontal cortex<sup>80</sup>. It has also been proposed that reduced alpha desynchronization in a network involving bilateral frontal and right-lateralized parietal regions may provide a specific measure of deficits in approach-related motivation in depression<sup>81</sup>. Using graph theoretical methods, another study shows that disrupted global and local network indices in MDD patients were revealed in alpha band<sup>11</sup>. In addition, features derived from alpha band have been revealed to be highly discriminating when distinguishing between MDD patients and healthy controls using binary classifiers<sup>32,34,37,82</sup>. These observations may be explained by the fact that alpha rhythm is thought to functionally inhibit cortical responses to unattended components of sensory input<sup>83</sup> and monitor important roles in both cognitive, social and emotion processing<sup>84,85</sup> that are closely altered in MDD. In contrast, a recent study show that resting state connectivity disruptions emerged mainly in beta frequency band (12.5–21 Hz), supporting the hypothesis that beta-band synchronization corresponds to a cognitive idling rhythm present in MDD<sup>21</sup>. In addition, a large study including 1344 participants showed increases in source EEG theta power across frontal regions of the brain<sup>86</sup>. This inconsistency in the results might be due to the difference in the subjects/patients characteristics as well the difference in methodological issues such as number of subjects, data pre-processing and data analysis (power spectra vs. connectivity for instance). Moreover, neural connectivity networks likely involve coordinated synchronizations across frequencies. Thus, extending the analysis to establish a model based on cross-frequency couplings for instance may be of interest.

**Limitations and methodological considerations.** First, the model was established using the first dataset composed of two groups (healthy and MDD) where the severity of depressive symptoms was assessed using BDI score. Although the BDI is a self-report questionnaire, as compared to the HDRS, was reported that BDI revealed high reliability and good correlation with measures of depression and anxiety<sup>87</sup>. However, like any self-report measure, BDI suffers from intrinsic limitations. Specifically, the score can be exaggerated or minimized based on the clinical environment<sup>88</sup>, fatigue<sup>89</sup>, and compromised cognitive functioning<sup>87</sup>.

Second, we are aware that subcortical regions are not easily accessible using scalp EEG, namely due to anatomical considerations. Unlike the layered cortex, a subcortical region would not have the necessary organization of pyramidal cells to give rise to localizable scalp-recorded EEG. Conversely, exploring subcortical regions may contribute to the modeling of depressive networks as many studies linked MDD with disruptions in amygdala<sup>7,18,90</sup>, a region centrally implicated in emotional and physiologic responses.

Third, a limitation of the current study is that one of the datasets included only healthy subjects and that some subjects experiencing mild depressive symptoms. We also believe that the correlation between observed and predicted HDRS scores needs to be validated in a larger sample (including MDD patients). In addition, longitudinal data (EEG data recorded from healthy subjects and MDD patients at several time points) is mandatory to validate and ultimately generalize the potential neuromarker.

Fourth, while the low-depressive network showed significant correlation between observed and predicted BDI scores in the Dataset 2, the external validation didn't reveal significant correlations between observed and predicted HDRS scores in Dataset 3. This observation suggests that the high-depressive network showed consistent prediction of depression severity across both datasets while the low-depressive network is less relevant to the prediction performance for Dataset 3. It is worth noting, however, that the Dataset 2 provided two separated group of subjects (healthy and MDD patients), while Dataset 3 didn't show a wide distribution of the HDRS score. More precisely, this latter mainly presents healthy subjects (with Hamilton scores spanning from 0 to 10) and some subjects experiencing mild depressive symptoms (Hamilton scores spanning from 10 to 14). This may explain the disability of the low-depressive network to discriminate between these subjects as it is expected to show similar low-network strength. Another explanatory possibility would be that normal sadness and pathological depression share the same continuum where major depression is defined based on the functional consequences of the symptoms (i.e. described as the pragmatic approach<sup>91,92</sup>). The continuum hypothesis between normal sadness and pathological major depression has also been supported at the symptom<sup>93</sup> and the cerebral levels<sup>94</sup> using data-driven latent analysis.

Additionally, we believe that a reliable interpretation of the predictive regression model is essential for clinical impact. One important issue that should be addressed is whether the significant correlation obtained between the predicted and the observed scores in Dataset 1 and 2 is biased by the two groups (HC and MDD) showing separated depression scores (see histograms in Figs. 3 and 4). To target this question, we have first assessed the predictive performance of the models not only using the correlation metric, but also using MAE and R-squared that are deemed to be more rigorous and sensitive to the scale of the prediction error. In addition, to quantify how much the prediction performances obtained are above/below chance, we employed a baseline regressor (also known as dummy regressor) that uses the average score of the training set as prediction. For the first dataset, results show that the established models (using high, low and both depressive networks as features) achieved MAE better than the error obtained by the dummy model, with no overlap between dummy MAE distributions (MAE = 1.18 ± 0.067) and MAE obtained by the model (MAE = 0.94 ± 0.06 using high depressive edges, MAE = 0.68 ± 0.09 using low depressive edges, MAE = 0.75 ± 0.091 when combining high and low depressive edges). For the second dataset, all but the model established using low depressive edges performed better than

the dummy revealing MAE = 1.73. To better assess if the correlations obtained are biased by the bimodal distributions, we performed a group-wise evaluation separately for MDD and HC. The correlation results (mean  $\pm$  std across folds) of Dataset 1 are reported in Table S1. For this dataset, results show that the group-wise correlations between the observed and predicted scores are considerable ( $> 0.35$ ) when both high and low depressive edges are used as predictive features. However, the use of high depressive edges solely results in acceptable correlations for the HC group but not for the MDD group. In contrast, the low depressive edges revealed considerable correlations in MDD group but not in HC group. A possible explanation is that the high (low) depressive edges are more related to the HC (MDD) group as their weights are positively (negatively) correlated with BDI scores. This issue highlights the importance of combining both high and low depressive networks as predictive features, especially in the bimodal case, as each predictive network (high/low) better describes a particular group. For dataset 2, results revealed that the predictive performance of the model separately for MDD is weak using each predictive network solely (Table S2). As in Dataset 1, the combination of both predictive edges (i.e. high and low) underlies a better performance of the model for both HC and MDD groups. This observation may be related to the difference in the data distribution between dataset 1 (showing HC subjects remarkably more than MDD patients) and dataset 2.

We should also note that the prediction error obtained in Dataset 3 (MAE = 0.47) is not sufficiently small taking into account the min–max range (1.6) and the interquartile (0.46) of the dataset. However, the permuted based p-value shows that the predictive performance is significant ( $p < 0.05$ ). In addition, the dummy mean model (returning the average of the training set as predicted score) resulted in MAE of 0.64, which suggests that the model performs better than expected by chance. On the other side, we are aware that the clinical needs require more robust predictions in terms of MAE, and R-squared.

Furthermore, internal validation was here assessed using a k-fold (fivefold and tenfold) cross validation procedure<sup>63,65</sup>. Although leave-one-out is the most popular cross validation method, it has been demonstrated that the k-fold cross validation provides more accurate prediction error. However, the choice of the number of folds (K) is critical and one should be prudent in K selection to ensure a good compromise between model bias and variance<sup>95–97</sup>. Referred to<sup>65</sup>, it is recommended to use 10 to 20% of data for the testing sample size. Thus, fivefold cross validation was used for the internal validation using the first dataset of 121 subjects so that 20% of data was used as testing sample in each iteration. Results of the internal validation using fivefold cross validation were also consistent with those obtained using a tenfold cross validation, suggesting that the main observations remain intact.

Previous studies showed that depression symptom severity strongly differ between elderly and young populations<sup>98</sup>. In the current study, we established the predictive model using resting-state functional connectivity acquired exclusively from young adults ( $< 40$  years). To better ensure that age is not correlated with the depression severity scores, we controlled for its confounding effects. It is noteworthy to mention that other demographic information (such as education background, years and drugs) should be, when available, collected and controlled in future studies.

In conclusion, our findings demonstrate the feasibility of resting-state EEG whole-brain connectome in predicting individual differences in depression severity, suggesting that it might serve as direct and non-invasive neuromarker that can ultimately support clinicians in a biologically-based characterization of MDD, which eventually will improve MDD prognostic and therapeutic decision.

Received: 18 August 2021; Accepted: 5 April 2022

Published online: 26 April 2022

## References

- Simon, S. A. & Nicolelis, M. A. L. Neurobiology of. *Depression* <https://doi.org/10.1093/bmb/lds004> (2011).
- Marcus, M., Yasamy, M. T., van Ommeren, M., & Chisholm, D. Depression, a global public health concern. *WHO Department of Mental Health and Substance Abuse* (2012).
- Cassano, P., & Fava, M. Depression and public health: An overview. [https://doi.org/10.1016/S0022-3999\(02\)00304-5](https://doi.org/10.1016/S0022-3999(02)00304-5) (2002).
- Insel, T. *et al.* Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *Am. J. Psychiatry* **167**(7), 748–751. <https://doi.org/10.1176/appi.ajp.2010.09091379> (2010).
- Schumann, G. *et al.* Stratified medicine for mental disorders. *Eur. Neuropsychopharmacol.* **24**(1), 5–50. <https://doi.org/10.1016/j.euroneuro.2013.09.010> (2014).
- Bullmore, E. & Sporns, O. Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci* **10**(3), 186–198. <https://doi.org/10.1038/nrn2575> (2009).
- Cullen, K. R. *et al.* Abnormal amygdala resting-state functional connectivity in adolescent depression. *JAMA Psychiat.* **71**(10), 1138–1147. <https://doi.org/10.1001/jamapsychiatry.2014.1087> (2014).
- Ye, M. *et al.* Changes of functional brain networks in major depressive disorder: A graph theoretical analysis of resting-state fMRI. *PLoS ONE* **10**(9), 1–16. <https://doi.org/10.1371/journal.pone.0133775> (2015).
- Sheline, Y. I., Price, J. L., Yan, Z., & Mintun, M. A. Resting-state functional MRI in depression unmasks increased connectivity between networks via the dorsal nexus. *Proc. Natl. Acad. Sci. U.S.A.* **107**(24), 11020–11025. <https://doi.org/10.1073/pnas.1000446107> (2010).
- Yu, Z. *et al.* Abnormal topology of brain functional networks in unipolar depression and bipolar disorder using optimal graph thresholding. *Prog. Neuro-Psychopharmacol. Biol. Psychiatry* **96**, 9758. <https://doi.org/10.1016/j.pnpbp.2019.109758> (2020).
- Shim, M., Im, C. H., Kim, Y. W. & Lee, S. H. Altered cortical functional network in major depressive disorder: A resting-state electroencephalogram study. *NeuroImage Clin.* **19**(1), 1000–1007. <https://doi.org/10.1016/j.nicl.2018.06.012> (2018).
- Damborská, A. *et al.* EEG resting-state large-scale brain network dynamics are related to depressive symptoms. *Front. Psych.* **10**(August), 1–10. <https://doi.org/10.3389/fpsy.2019.00548> (2019).
- Greicius, M. D. *et al.* Resting-state functional connectivity in major depression: abnormally increased contributions from subgenual cingulate cortex and thalamus. *Biol. Psychiat.* **62**(5), 429–437. <https://doi.org/10.1016/j.biopsych.2006.09.020> (2007).

14. Anand, A., Li, Y., Wang, Y., Lowe, M. J. & Dzemidzic, M. Resting state corticolimbic connectivity abnormalities in unmedicated bipolar disorder and unipolar depression. *Psychiatry Res. Neuroimaging* <https://doi.org/10.1016/j.psychres.2008.03.012> (2009).
15. Dutta, A., McKie, S. & Deakin, J. F. W. Resting state networks in major depressive disorder. *Psychiatry Res. Neuroimaging* <https://doi.org/10.1016/j.psychres.2014.10.003> (2014).
16. Hou, Z. *et al.* Divergent topological architecture of the default mode network as a pretreatment predictor of early antidepressant response in major depressive disorder. *Sci. Rep.* <https://doi.org/10.1038/srep39243> (2016).
17. Sheline, Y. I., Price, J. L., Yan, Z. & Mintun, M. A. Resting-state functional MRI in depression unmasks increased connectivity between networks via the dorsal nexus. *Proc. Natl. Acad. Sci. U.S.A.* <https://doi.org/10.1073/pnas.1000446107> (2010).
18. Connolly, C. G. *et al.* Resting-state functional connectivity of subgenual anterior cingulate cortex in depressed adolescents. *Biol. Psychiat.* <https://doi.org/10.1016/j.biopsych.2013.05.036> (2013).
19. Zhou, Y. *et al.* Early altered resting-state functional connectivity predicts the severity of post-traumatic stress disorder symptoms in acutely traumatized subjects. *PLoS ONE* <https://doi.org/10.1371/journal.pone.0046833> (2012).
20. Wang, X., Öngür, D., Auerbach, R. P. & Yao, S. Cognitive vulnerability to major depression: View from the intrinsic network and cross-network interactions. *Harv. Rev. Psychiatry* **24**(3), 188–201. <https://doi.org/10.1097/HRP.0000000000000081> (2016).
21. Albert, K. M., Potter, G. G., Boyd, B. D., Kang, H. & Taylor, W. D. Brain network functional connectivity and cognitive performance in major depressive disorder. *J. Psychiatr. Res.* **110**, 51–56. <https://doi.org/10.1016/j.jpsychires.2018.11.020> (2019).
22. Bi, K. *et al.* Abnormal early dynamic individual patterns of functional networks in low gamma band for depression recognition. *J. Affect. Disord.* **238**, 366–374. <https://doi.org/10.1016/j.jad.2018.05.078> (2018).
23. Whitton, A. E. *et al.* Electroencephalography source functional connectivity reveals abnormal high-frequency communication among large-scale functional networks in depression. *Biol. Psychiatry Cognit. Neurosci. Neuroimaging* **3**(1), 50–58. <https://doi.org/10.1016/j.bpsc.2017.07.001> (2018).
24. Wei, M. *et al.* Abnormal dynamic community structure of the salience network in depression. *J. Magn. Reson. Imaging* <https://doi.org/10.1002/jmri.25429> (2017).
25. Chen, T. *et al.* Anomalous single-subject based morphological cortical networks in drug-naive, first-episode major depressive disorder. *Hum. Brain Mapp.* <https://doi.org/10.1002/hbm.23534> (2017).
26. Greicius, M. D. *et al.* Resting-state functional connectivity in major depression: Abnormally increased contributions from subgenual cingulate cortex and thalamus. *Biol. Psychiat.* <https://doi.org/10.1016/j.biopsych.2006.09.020> (2007).
27. Jiang, X. *et al.* Connectome analysis of functional and structural hemispheric brain networks in major depressive disorder. *Transl. Psychiatry* **9**(1), 1. <https://doi.org/10.1038/s41398-019-0467-9> (2019).
28. Kaiser, R. H., Andrews-Hanna, J. R., Wager, T. D. & Pizzagalli, D. A. Large-scale network dysfunction in major depressive disorder: A meta-analysis of resting-state functional connectivity. *JAMA Psychiat.* **72**(6), 603–611 (2015).
29. Broyd, S. J. *et al.* Default-mode brain dysfunction in mental disorders: A systematic review. *Neurosci. Biobehav. Rev.* **33**(3), 279–296. <https://doi.org/10.1016/j.neubiorev.2008.09.002> (2009).
30. Gusnard, D. A., Akbudak, E., Shulman, G. L. & Raichle, M. E. Medial prefrontal cortex and self-referential mental activity: Relation to a default mode of brain function. *PNAS* **98**(7), 4259–4264. <https://doi.org/10.1073/pnas.071043098> (2001).
31. Thoduparambil, P. P., Dominic, A. & Varghese, S. M. EEG-based deep learning model for the automatic detection of clinical depression. *Phys. Eng. Sci. Med.* <https://doi.org/10.1007/s13246-020-00938-4> (2020).
32. Uyulan, C. *et al.* Major depressive disorder classification based on different convolutional neural network models: Deep learning approach. *Clin. EEG Neurosci.* <https://doi.org/10.1177/1550059420916634> (2020).
33. Xie, Y. *et al.* Anxiety and depression diagnosis method based on brain networks and convolutional neural networks. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* **2020**, 1503–1506. <https://doi.org/10.1109/EMBC44109.2020.9176471> (2020).
34. Rolle, C. E. *et al.* Cortical connectivity moderators of antidepressant vs placebo treatment response in major depressive disorder: Secondary analysis of a randomized clinical trial. *JAMA Psychiat.* **77**(4), 397–408. <https://doi.org/10.1001/jamapsychiatry.2019.3867> (2020).
35. Zhdanov, A. *et al.* Use of machine learning for predicting escitalopram treatment outcome from electroencephalography recordings in adult patients with depression. *JAMA Netw. Open* **3**(1), 8377. <https://doi.org/10.1001/jamanetworkopen.2019.18377> (2020).
36. Mahato, S., Goyal, N., Ram, D. & Paul, S. Detection of depression and scaling of severity using six channel EEG data. *J. Med. Syst.* **44**(7), 118. <https://doi.org/10.1007/s10916-020-01573-y> (2020).
37. Mohammadi, Y. & Moradi, M. H. Prediction of depression severity scores based on functional connectivity and complexity of the EEG signal. *Clin. EEG Neurosci.* <https://doi.org/10.1177/1550059420965431> (2020).
38. Trambaiolli, L. R. & Biazoli, C. E. Resting-state global EEG connectivity predicts depression and anxiety severity. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* **2020**, 3707–3710. <https://doi.org/10.1109/EMBC44109.2020.9176161> (2020).
39. Jollans, L. *et al.* Quantifying performance of machine learning methods for neuroimaging data. *Neuroimage* **199**, 351–365. <https://doi.org/10.1016/j.neuroimage.2019.05.082> (2019).
40. Jaworska, N., Blier, P., Fusee, W. & Knott, V. Alpha power, alpha asymmetry and anterior cingulate cortex activity in depressed males and females. *J. Psychiatr. Res.* <https://doi.org/10.1016/j.jpsychires.2012.08.003> (2012).
41. Knott, V., Mahoney, C., Kennedy, S. & Evans, K. EEG power, frequency, asymmetry and coherence in male depression. *Psychiatry Res. Neuroimaging* [https://doi.org/10.1016/S0925-4927\(00\)00080-9](https://doi.org/10.1016/S0925-4927(00)00080-9) (2001).
42. Fingelkurts, A. A. *et al.* Impaired functional connectivity at EEG alpha and theta frequency bands in major depression. *Hum. Brain Mapp.* <https://doi.org/10.1002/hbm.20275> (2007).
43. Hassan, M. & Wendling, F. Electroencephalography source connectivity. *IEEE Signal Process. Mag.* **1**, 81–96. <https://doi.org/10.1109/MSP.2017.2777518> (2018).
44. Shen, X. *et al.* Using connectome-based predictive modeling to predict individual behavior from brain connectivity. *Nat. Protoc.* <https://doi.org/10.1038/nprot.2016.178> (2017).
45. Cavanagh, J. F., Bismark, A. W., Frank, M. J. & Allen, J. J. B. Multiple dissociations between comorbid depression and anxiety on reward and punishment processing: Evidence from computationally informed EEG. *Comput. Psychiatry* [https://doi.org/10.1162/cpsy\\_a\\_00024](https://doi.org/10.1162/cpsy_a_00024) (2019).
46. Kabbara, A. *et al.* Reduced integration and improved segregation of functional brain networks in Alzheimer's disease. *J. Neural Eng.* **1**, 1 (2018).
47. Rizkallah, J. *et al.* Dynamic reshaping of functional brain networks during visual object recognition. *J. Neural Eng.* <https://doi.org/10.1088/1741-2552/aad7b1> (2018).
48. Kabbara, A., Falou, W. E. L., Khalil, M., Wendling, F. & Hassan, M. The dynamic functional core network of the human brain at rest. *Sci. Rep.* **7**(1), 2936 (2017).
49. Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D. & Leahy, R. M. Brainstorm: A user-friendly application for MEG/EEG analysis. *Comput. Intell. Neurosci.* <https://doi.org/10.1155/2011/879716> (2011).
50. Babayan, A. *et al.* Data descriptor: A mind-brain-body dataset of MRI, EEG, cognition, emotion, and peripheral physiology in young and old adults. *Sci. Data* <https://doi.org/10.1038/sdata.2018.308> (2019).
51. Hamilton, M. A rating scale for depression. *J. Neurol. Neurosurg. Psychiatry* <https://doi.org/10.1136/jnnp.23.1.56> (1960).
52. Hassan, M. & Wendling, F. Electroencephalography source connectivity: toward high time/space resolution brain networks. *IEEE Signal Process. Mag.* **1**, 1–25 (2018).



53. Gramfort, A., Papadopoulo, T., Olivi, E. & Clerc, M. OpenMEEG: Opensource software for quasistatic bioelectromagnetics. *Biomed. Eng. Online* **9**, 45. <https://doi.org/10.1186/1475-925X-9-45> (2010).
54. Desikan, R. S. *et al.* An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* **31**, 968–980. <https://doi.org/10.1016/j.neuroimage.2006.01.021> (2006).
55. Hamalainen, M. S. & Ilmoniemi, R. J. Interpreting magnetic fields of the brain: Minimum norm estimates. *Med. Biol. Eng. Compu.* **32**(1), 35–42. <https://doi.org/10.1007/BF02512476> (1994).
56. Mheich, A. *et al.* HD-EEG for tracking sub-second brain dynamics during cognitive tasks. *Sci. Data* **8**(1), 1. <https://doi.org/10.1038/s41597-021-00821-1> (2021).
57. Hassan, M. *et al.* Identification of interictal epileptic networks from dense-EEG. *Brain Topogr.* **30**(1), 60–76. <https://doi.org/10.1007/s10548-016-0517-z> (2017).
58. Wang, Z., Goerlich, K. S., Ai, H., Aleman, A., Luo, Y., & Xu, P. Connectome-based predictive modeling of individual anxiety. *bioRxiv*, 2020.01.30.926980. <https://doi.org/10.1101/2020.01.30.926980> (2020).
59. Giancardo, L., *et al.* Longitudinal connectome-based predictive modeling for REM sleep behavior disorder from structural brain connectivity. <https://doi.org/10.1117/12.2293835> (2018).
60. Beaty, R. E. *et al.* Robust prediction of individual creative ability from brain functional connectivity. *Proc. Natl. Acad. Sci. U.S.A.* <https://doi.org/10.1073/pnas.1713532115> (2018).
61. Yoo, K. *et al.* Connectome-based predictive modeling of attention: Comparing different functional connectivity features and prediction methods across datasets. *Neuroimage* <https://doi.org/10.1016/j.neuroimage.2017.11.010> (2018).
62. Feng, C., Wang, L., Li, T. & Xu, P. Connectome-based individualized prediction of loneliness. *Social Cognit. Affect. Neurosci.* <https://doi.org/10.1093/scan/nsz020> (2019).
63. Scheinost, D. *et al.* Ten simple rules for predictive modeling of individual differences in neuroimaging. *Neuroimage* **193**, 35–45. <https://doi.org/10.1016/j.neuroimage.2019.02.057> (2019).
64. Rosenberg, M. D. *et al.* A neuromarker of sustained attention from whole-brain functional connectivity. *Nat. Neurosci.* **19**(1), 165–171. <https://doi.org/10.1038/nn.4179> (2016).
65. Poldrack, R. A., Huckins, G. & Varoquaux, G. Establishment of best practices for evidence for prediction: A review. *JAMA Psychiat.* **77**(5), 534–540. <https://doi.org/10.1001/jamapsychiatry.2019.3671> (2020).
66. Feng, C. *et al.* Individualized prediction of trait narcissism from whole-brain resting-state functional connectivity. *Hum. Brain Mapp* **39**(9), 3701–3712. <https://doi.org/10.1002/hbm.24205> (2018).
67. Awad, M., & Khanna, R. Support vector regression. in *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*, Awad, M., & Khanna, R., Eds. Berkeley, CA: Apress, pp. 67–80. [https://doi.org/10.1007/978-1-4302-5990-9\\_4](https://doi.org/10.1007/978-1-4302-5990-9_4) (2015).
68. Sestieri, C., Corbetta, M., Romani, G. L. & Shulman, G. L. Episodic memory retrieval, parietal cortex, and the default mode network: Functional and topographic analyses. *J. Neurosci.* <https://doi.org/10.1523/JNEUROSCI.3335-10.2011> (2011).
69. Cooney, R. E., Joormann, J., Eugène, F., Dennis, E. L. & Gotlib, I. H. Neural correlates of rumination in depression. *Cognit. Affect. Behav. Neurosci.* <https://doi.org/10.3758/CABN.10.4.470> (2010).
70. Monkul, E. S. *et al.* Abnormal resting state corticolimbic blood flow in depressed unmedicated patients with major depression: A 15O-H 2O PET study. *Hum. Brain Mapp.* <https://doi.org/10.1002/hbm.21212> (2012).
71. Phillips, M. L., Drevets, W. C., Rauch, S. L. & Lane, R. Neurobiology of emotion perception II: Implications for major psychiatric disorders. *Biol. Psychiat.* [https://doi.org/10.1016/S0006-3223\(03\)00171-9](https://doi.org/10.1016/S0006-3223(03)00171-9) (2003).
72. Pinel, J. P. J. Biopsychology 5th edition. *Pearson. Education* (2003).
73. Hu, H., Cui, Y. & Yang, Y. Circuits and functions of the lateral habenula in health and in disease. *Nat. Rev. Neurosci.* **21**(5), 277–295. <https://doi.org/10.1038/s41583-020-0292-4> (2020).
74. Yang, Y., Wang, H., Hu, J. & Hu, H. Lateral habenula in the pathophysiology of depression. *Curr. Opin. Neurobiol.* **48**, 90–96. <https://doi.org/10.1016/j.conb.2017.10.024> (2018).
75. Browne, C. A., Hammack, R. & Lucki, I. Dysregulation of the lateral habenula in major depressive disorder. *Front Synaptic Neurosci.* **10**, 46. <https://doi.org/10.3389/fnsyn.2018.00046> (2018).
76. Covington, H. E. *et al.* Antidepressant effect of optogenetic stimulation of the medial prefrontal cortex. *J. Neurosci.* **30**(48), 16082–16090. <https://doi.org/10.1523/JNEUROSCI.1731-10.2010> (2010).
77. Iacoboni, M. & Zaidel, E. Interhemispheric visuo-motor integration in humans: The role of the superior parietal cortex. *Neuropsychologia* <https://doi.org/10.1016/j.neuropsychologia.2003.10.007> (2004).
78. Ricardo-Garcell, J. *et al.* EEG sources in a group of patients with major depressive disorders. *Int. J. Psychophysiol.* **71**(1), 70–74. <https://doi.org/10.1016/j.ijpsycho.2008.07.021> (2009).
79. van der Vinne, N., Vollebregt, M. A., van Putten, M. A. J. M. & Arns, M. Frontal alpha asymmetry as a diagnostic marker in depression: Fact or fiction: A meta-analysis. *NeuroImage Clin.* **1**, 1. <https://doi.org/10.1016/j.nicl.2017.07.006> (2017).
80. Olbrich, S., Tränkner, A., Chittka, T., Hegerl, U. & Schönknecht, P. Functional connectivity in major depression: Increased phase synchronization between frontal cortical EEG-source estimates. *Psychiatry Res. Neuroimaging* <https://doi.org/10.1016/j.pscychres.2014.02.010> (2014).
81. Messerotti Benvenuti, S., Buodo, G., Mennella, R., Dal Bò, E. & Palomba, D. Appetitive and aversive motivation in depression: The temporal dynamics of task-elicited asymmetries in alpha oscillations. *Sci. Rep.* **9**(1), 7129. <https://doi.org/10.1038/s41598-019-53639-8> (2019).
82. Trambaiolli, L. R., & Biazoli, C.E. Resting-state global EEG connectivity predicts depression and anxiety severity. in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Montreal, QC, Canada, pp. 3707–3710 <https://doi.org/10.1109/EMBC44109.2020.9176161> (2020).
83. Desimone, R. & Duncan, J. Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* **18**, 193–222. <https://doi.org/10.1146/annurev.ne.18.030195.001205> (1995).
84. Aftanas, L. I. & Golocheikine, S. A. Human anterior and frontal midline theta and lower alpha reflect emotionally positive state and internalized attention: High-resolution EEG investigation of meditation. *Neurosci. Lett.* [https://doi.org/10.1016/S0304-3940\(01\)02094-8](https://doi.org/10.1016/S0304-3940(01)02094-8) (2001).
85. Aftanas, L. I., Varlamov, A. A., Pavlov, S. V., Makhnev, V. P. & Reva, N. V. Time-dependent cortical asymmetries induced by emotional arousal: EEG analysis of event-related synchronization and desynchronization in individually defined frequency bands. *Int. J. Psychophysiol.* [https://doi.org/10.1016/S0167-8760\(01\)00194-5](https://doi.org/10.1016/S0167-8760(01)00194-5) (2002).
86. Arns, M. *et al.* Frontal and rostral anterior cingulate (rACC) theta EEG in depression: Implications for treatment outcome?. *Eur. Neuropsychopharmacol.* <https://doi.org/10.1016/j.euroneuro.2015.03.007> (2015).
87. Wang, Y. P. & Gorenstein, C. Assessment of depression in medical patients: A systematic review of the utility of the Beck Depression Inventory-II. *Clinics* [https://doi.org/10.6061/clinics/2013\(09\)15](https://doi.org/10.6061/clinics/2013(09)15) (2013).
88. Bowling, A. Mode of questionnaire administration can have serious effects on data quality. *J. Public Health* <https://doi.org/10.1093/pubmed/fdi031> (2005).
89. Moore, M. J., Moore, P. B., & Shaw, P. J. Mood disturbances in motor neurone disease. [https://doi.org/10.1016/S0022-510X\(98\)00203-2](https://doi.org/10.1016/S0022-510X(98)00203-2) (1998).
90. Price, J. L. & Drevets, W. C. Neurocircuitry of mood disorders. *Neuropsychopharmacology* <https://doi.org/10.1038/npp.2009.104> (2010).

91. Maj, M. When does depression become a mental disorder?. *Br. J. Psychiatry* **199**(2), 85–86. <https://doi.org/10.1192/bjp.bp.110.089094> (2011).
92. M. Maj (2016) The continuum of depressive states in the population and the differential diagnosis between ‘normal’ sadness and clinical depression. in *Sadness or Depression? International Perspectives on the Depression Epidemic and Its Meaning*, J. C. Wakefield and S. Demazeux, Eds. Dordrecht: Springer Netherlands, 2016, pp. 29–38. [https://doi.org/10.1007/978-94-017-7423-9\\_3](https://doi.org/10.1007/978-94-017-7423-9_3).
93. Ruscio, J. & Ruscio, A. M. Informing the continuity controversy: A taxometric analysis of depression. *J. Abnorm. Psychol.* **109**(3), 473–487 (2000).
94. Ing, A. *et al.* Identification of neurobehavioural symptom groups based on shared brain mechanisms. *Nat. Hum. Behav.* **3**(12), 1306–1318. <https://doi.org/10.1038/s41562-019-0738-8> (2019).
95. Breiman, L. & Spector, P. Submodel selection and evaluation in regression. The X-Random Case. *Int. Stat. Rev.* **60**(3), 291–319. <https://doi.org/10.2307/1403680> (1992).
96. Pereira, F., Mitchell, T. & Botvinick, M. Machine learning classifiers and fMRI: A tutorial overview. *Neuroimage* **45**(1), S199–S209. <https://doi.org/10.1016/j.neuroimage.2008.11.007> (2009).
97. Varoquaux, G. *et al.* Assessing and tuning brain decoders: Cross-validation, caveats, and guidelines. *Neuroimage* **145**, 166–179. <https://doi.org/10.1016/j.neuroimage.2016.10.038> (2017).
98. Ko, S. M., Kua, E. H. & Chow, M. H. Depression of young and elderly patients. *Singapore Med. J.* **38**(10), 439–441 (1997).

## Acknowledgements

This work was financed by the Rennes University, the Institute of Clinical Neuroscience of Rennes (Projects named EEGCog and EEGNET3). This work was also financed by the AZM and SAADE Association, Tripoli, Lebanon and by the National Council for Scientific Research (CNRS) in Lebanon. Authors would like to thank Campus France, Programme Hubert Curien CEDRE (PROJET N° 42257YA) and the Lebanese Association for Scientific Research (LASER) for their support.

## Author contributions

A.K. and M.H. analysed the data. A.K., M.H., G.R. and P.B. wrote the main manuscript text. All authors interpret the data and reviewed the manuscript.

## Funding

This article was funded by Université Libanaise, Institute of Clinical Neuroscience of Rennes.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-10949-8>.

**Correspondence** and requests for materials should be addressed to M.H.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022